Journal of Algebra **221,** 475–496 (1999)
Article ID jabr.1999.7922, available online at http://www.idealibrary.com on IDEAL®

Which Words Spell "Almost Nilpotent?"

Sarah Black

Institute of Mathematics, Hebrew University, Jerusalem 91904, Israel

Communicated by Efim Zelmanov

Received June 5, 1998

A class of words is introduced with the property that a finite group satisfying an identity of length N defined by a member of this class is nilpotent-of-N-bounded-class-by-N-bounded-exponent. Conversely, we see that this class of words *characterises* such groups; a group which is nilpotent-of-class-c-by-exponent-at-most-c, satisfies an identity from this class whose length is c-bounded. Related results are discussed.

INTRODUCTION

The aim of the present paper is as the title suggests: given a word w on an alphabet $\mathscr{X}$, where $\mathscr{X} \subseteq \mathscr{F}$, the free group of infinite rank, we should like to develop criteria to determine when the variety generated by *finite* groups defined by w is nilpotent-of-bounded-class-by-bounded-exponent, for uniform such bounds which are functions of the length of the word alone.

Some Important Points to Note

(1) *Recall that a positive word* in an alphabet $\mathscr{X}$ is a word u in $\mathscr{X}$ not involving inverses. Similarly, a *positive law* is a law of the form $u \equiv v$ for u, v positive words. Let us call $w = w(\mathscr{X})$ a word which *yields* a positive law, if $w \equiv uv^{-1}$ for u, v positive words in $\mathscr{X}$. Results of Semple and Shalev [8; 9, Theorem B], in conjunction with [3, Theorem 1], imply that there exists a function $f = f(N)$, so that if w is a word of length N which yields a positive law, then for every (residually) finite group G satisfying the law $w \equiv 1$, we have $\gamma_{f+1}(G^f) = 1$; thus such a G is nilpotent-of-class $\leq f$-by-exponent $\leq f$. Thus this paper extends these results.

0021-8693/99 $30.00

It turns out that the words we tackle are also generalizations of the well-known Engel identities and of the Milnor identities introduced by Point [7]. We also mention a related condition introduced by Kim and Rhemtulla [5]; namely, a group is said to be *strongly restrained* if there exists an integer n such that $\langle x^{\langle y \rangle} \rangle$ can be generated by n elements for all x, y in G. It turns out that finitely generated (residually) finite groups satisfying our criteria are strongly restrained by virtue of their being nilpotent-of-bounded-class-by-bounded-index (see [5, Theorem A]). On the other hand, it would be interesting to be able to derive a direct proof of this fact.

(2) It is abundantly clear that certain words can in no way serve to define such "nilpotent-by-bounded-exponent" varieties; for example, any word $w \in \mathscr{F}^{(d)}$, the dth derived subgroup of $\mathscr{F}$, for $d \geq 2$. Indeed, words of this type will be implicitly precluded from our study (see Note (4) after the theorem in Sect. I).

(3) As a curiosity, it is worth mentioning that our results, among other things, enable us to prove that, for example, the following word defines a nilpotent-by-bounded-exponent-variety-generated-by-finite-groups; a fact which is hardly obvious at the outset!

$$w(x, y) \equiv x^{-2}y^{-2}x^{2}y^{2}x^{-10}y^{-2}x^{10}y^{2}x^{-2}y^{-4}x^{-6}y^{6}x^{-2}y^{-2}x^{10}y^{-4}x^{-8}y^{6}x^{8}y^{-2}.$$

Outline of the Paper

In Section I we present a verbal criterion for a given variety generated by finite groups to be a nilpotent-by-bounded-exponent variety, in the sense suggested above. However, since the result is not, on the face of it, an *effective* criterion for all words, we devote the second section to a presentation of some methods which can be used to determine when a given word satisfies the criterion. In Section III we present some consequences of the theory developed in the first two sections.

This work is replete with examples, whose aim is to demonstrate these techniques, as well as their limitations. While the sufficient conditions which are developed in Section II do not cover all words, it turns out that words to which they cannot be applied have certain very uniform and specified structures.

We thus end with some open questions, based on this on the one hand, and on the basis of some counter-examples of this type, on the other.

Note. Some of the results in Section I have been obtained independently by Point for a wider class of groups, the class of elementary amenable groups (see [7]); however, our result is proved for a more general class of words than the Milnor identities introduced in that paper, and also makes use of the work of [3] to yield sharper statements in each case.

SECTION I

We first introduce some definitions which we shall require for the statement of the theorem.

Since, by the Nielsen−Schreier theorem, subgroups of free groups are free, we let $F_2 := F_2(x, y) \subseteq \mathscr{F}$ denote the free group of rank 2 with generators x, y viewed as a subgroup of $\mathscr{F}$, let $X := \langle x \rangle$, the cyclic group generated by x, $Y := \langle y \rangle$, X^Y (resp. Y^X) the normal closure of X in F_2 (resp. the normal closure of Y in F_2).

Clearly $X^Y = \langle x \rangle^{\langle y \rangle}$, hence the choice of notation. Denote, as usual, in a group G, by G' its first derived subgroup, by $G^{(d)}$, $d = 2, \dots$ its dth derived subgroup, and by $\gamma_i(G)$ the ith member of its lower central series, and set $G^m := \langle g^m | g \in G \rangle$. Note that since $X^Y = X \cdot F_2'$, $Y^X = Y \cdot F_2'$, we have $X^Y \cap Y^X = F_2'$.

Now suppose that $w(x, y) = x^{a_1} y^{b_1} x^{a_2} \cdots x^{a_n} y^{b_n}$ is a word in F_2, with $\sum_{i=1}^n a_i = \sum_{j=1}^n b_j = 0$, and $\sum_{i=1}^n |a_i| + \sum_{j=1}^n |b_j| = N$. We call such a word w a *homogeneous word of length N in* $\{x, y\}$.

Note that by homogeneity we may rewrite w in the form

$$w(x, y) = x^{a_1} (x^{a_2})^{y^{s_2}} (x^{a_3})^{y^{s_3}} \dots (x^{a_n})^{y^{s_n}}, \tag{1}$$

where $x^y := y^{-1}xy$, $s_i = -\sum_{j=1}^{i-1} b_j$, $(i = 1, \dots, n)$, or, in the form

$$w(x, y) = (y^{b_1})^{x^{u_1}} (y^{b_2})^{x^{u_2}} \dots (y^{b_n}), \tag{2}$$

where $u_i = -\sum_{j=1}^{i} a_j$ $(i = 1, \dots, n)$. (Note that $s_1 = u_n = 0$ by the homogeneity of w, and are thus omitted.)

Since clearly every commutator word is homogeneous, it is clear by the above remarks that w is a homogeneous word iff w is a commutator word, i.e., iff $w \in F_2'$. (The fact that a homogeneous word is in F_2' can in fact be seen directly, by noting that its image in F_2/F_2' is $x^{\sum_{i=1}^n a_i} y^{\sum_{j=1}^n b_j} \equiv 1$.)

The following discussion will motivate the definition of an *efficient* commutator word.

Consider form (1) for w; thus, consider $w \in X^Y$. "Collect together" those terms $(x^{a_i})^{y^{s_j}}, (x^{a_\ell})^{y^{s_\ell}}$ for which $s_j = s_\ell$, by "commutator collection," subsequently writing it as $(x^{a_i + a_\ell})^{y^{s_j}}$; that is, we shall group together all such expressions which are conjugated by the same power of y, by performing commutator collection modulo $(X^Y)'$, and write the commutator "remainder" in $(X^Y)'$ as c (see form (3)). We thus obtain a word

$$w(x, y) \equiv x^{t_1} (x^{t_2})^{y^{s_2}} (x^{t_3})^{y^{s_3}} \cdots (x^{t_m})^{y^{s_m}} c, \tag{3}$$

where $c \in (X^Y)'$, the $s_1, \dots, s_m$ for $m \le n$ are all distinct, and the t_j respectively represent the non-zero sums of the exponents a_j of those x^{a_j}

which are conjugated by the same power of y, where, if such a sum equals zero, we omit the term from the expression. A similar procedure for w in form (2) $\in Y^X$ renders

$$w(x,y) \equiv (y^{v_1})^{x^{u_1}}(y^{v_2})^{x^{u_2}}\dots(y^{v_r})^{x^{u_r}}d, \tag{4}$$

where $d \in (Y^X)'$, $r \le n$, the u_j for $j = 1, \dots, r$ are mutually distinct, and $v_i \ne 0$, $i = 1, \dots, r$.

We illustrate this procedure with an example;

EXAMPLE 1. $w_1(x,y) \equiv x^{-2}y^{-2}xy^2x^2y^{-3}x^{-1}y^3$ is indeed a commutator word, and takes the forms

(1) $w_1(x,y) \equiv x^{-2}(x)^{y^2}x^2(x^{-1})^{y^3}$
(2) $w_1(x,y) \equiv (y^{-2})^{x^2}(y^2)^x(y^{-3})^{x^{-1}}y^3$
(3) $w_1(x,y) \equiv (x)^{y^2}(x^{-1})y^3c$
(4) $w_1(x,y) \equiv (y^{-2})^{x^2}(y^2)^x(y^{-3})^{x^{-1}}y^3$.

Note that form (4) is identical to form (2) by virtue of the fact that all the conjugating powers of x in form (2) are distinct, so there is, in fact, no "collection" to be done.

DEFINITION. Let $w = w(x,y)$ be a homogeneous word in $\{x, y\}$. Call w an *efficient*[1] word if

(1) $w \ne (X^Y)' \cap (Y^X)'$ and
(2) $GCD(\{t_i\}_{1 \le i \le m}, \{v_j\}_{1 \le j \le r}) = 1$.

EXAMPLES OF EFFICIENT AND NONEFFICIENT WORDS. (1) Note that w_1 above is efficient.

(2) $w_2 := [x, {}_n y]$ the well-known Engel identity, is an efficient word, for, when written in form (3), we have $w_2 \equiv x^{(y-1)^n}c$, where $(y-1)^n$ is expanded as a binomial of degree n, giving

$$w_2 \equiv \prod_{i=0}^{n} x^{\binom{n}{i}(-1)^i y^{n-i}}c \equiv \prod_{i=0}^{n}\left(x^{\binom{n}{i}(-1)^i}\right)^{y^{n-i}}c,$$

and so w_2 is efficient. (Note that indeed the exponent sum of the x's occurring in this form equals $\sum_{i=0}^{n}\binom{n}{i}(-1)^i = (1-1)^n = 0$ as required.)

[1] The term alludes to the fact that a finite group satisfying a *law* $w \equiv 1$ for such words w, will be seen to have bounds on the ranks of *all* its chief factors.

Notice two families of *non*efficient words, to which we shall subsequently return,

(3) $w_{3,n_0} = [x^{n_0}, y^{n_0}]$ and

(4) $w_{4,n_0} = [x, y]^{n_0}$,

where n_0 is any fixed natural number greater than 1. (They will be seen to represent, in general, a fundamental difference between, on the one hand, a word $w(x, y)$ for which $w(x^{n_0}, y^{n_0})$ is efficient *in the arguments* x^{n_0} *and* y^{n_0}, and on the other hand, a word $w(x, y) = v(x, y)^{n_0}$ for which $v(x, y)$ is efficient.)

Let G be a group and $w = w(\mathscr{X})$ a word on an alphabet $\mathscr{X} = \{x_1, \ldots, x_n\}$. We say G satisfies the *group law* $w \equiv 1$ if, for all subsets $S = \{g_1, \ldots, g_n\} \subseteq G$ (including those for which the g_i's are not necessarily mutually distinct), we have $w(S) \equiv 1$. Similarly, G satisfies an *efficient law* if G satisfies the group law $w \equiv 1$, where w is efficient. Let $\mathscr{X}, \mathscr{Y} \subseteq \mathscr{F}$.

Let $w = w(\mathscr{X})$ be a word in $\mathscr{X}$, which, without loss of generality, we assume to be finite, say, $\mathscr{X} = \{x_1, \ldots, x_n\}$; let $\mathscr{Y} = \{y_1, \ldots, y_m\}$.

We say that $\phi = \phi(\mathscr{Y})$ is a *result of* w if there exist substitutions

$$\alpha : \mathscr{X} \to \mathscr{F}(\mathscr{Y})$$

$$x_i \mapsto \alpha(x_i) \quad i = 1, \ldots, n$$

with

$$\phi(\mathscr{Y}) := w\big((\alpha(x_i)_{i=1,\ldots n})\big).$$

In other words, ϕ is a result of w if ϕ is a homomorphic image of w for some homomorphism α which is defined on the generators $\mathscr{X}$ and extended naturally to w by multiplicativity. Note that if ϕ is a result of w, then ϕ is a group law whenever w is. If w is also a result of ϕ, we say that w and ϕ are *equivalent*.

Denote, for each i, by $\ell_{\mathscr{Y}}(\alpha(x_i)) :=$ minimal length of a word in the elements of $\mathscr{Y}$ representing $\alpha(x_i)$.

We are now ready to state the theorem.

THEOREM. (i) *There exists a function $f = f(N)$ so that, if $w = w(\mathscr{X})$ is any efficient word of length N in an alphabet $\mathscr{X}$, then any (residually) finite group G satisfying w is nilpotent-of-class-at-most-f-by-exponent-at-most-f.*

(ii) *Let w have a result $\phi = \phi(c, d)$ which is efficient. Then the same conclusions hold as in* (i), *but with the nilpotency class and the exponent bounded in terms of* $\max_i(\ell_{\{c,d\}}(\alpha(x_i)))$ *as well as f.*
Conversely,

(iii) *If G is any group (not necessarily finite), satisfying $\gamma_{c+1}(G^c) = 1$, then G satisfies an efficient law, whose length is c-bounded.*

Clearly, proving parts (i) and (ii) for finite groups equivalently justifies their veracity for residually finite groups, as well as for the variety generated by all finite groups satisfying the given law (see, for example, the Corollary below).

COROLLARY. *Let $\mathscr{V}$ be a variety generated by finite groups. Then* $\mathscr{V} \subseteq \mathscr{N}_{c+1}\mathscr{B}_e \Leftrightarrow$ *there exists an efficient word w of c, e bounded length, which is a law in $\mathscr{V}$.*

(Here $\mathscr{N}_{c+1}$ denotes, as usual, the variety of nilpotent groups of class c and $\mathscr{B}_e$ the variety of locally finite groups of exponent e.)

Notes. (1) The theorem can be seen to be a generalization of Shalev's result on positive laws [9, Theorem B] as well as a result from [3, Theorem 1]. Indeed, we shall show in the next section that any positive law of length N has an efficient result of N-bounded length.

(2) Note that Engel words are efficient (see Example 2 above). Thus the theorem is reminiscent of a result of Wilson [10, Theorem 2] that a finitely generated residually finite group satisfying an Engel law is *nilpotent*. Indeed, while Wilson's result can be derived from the theorem, it can also be derived from results of Zelmanov [12, 13] which, in turn, are used implicitly in the proof of our theorem.

(3) It would be interesting to be able to strengthen the theorem to state that "w defines a nilpotent-of-bounded-class-by-bounded-exponent variety generated by finite groups $\Leftrightarrow$ w has a *result* which is efficient."

(4) Suppose, as mentioned in the Introduction, that $w \in F_2^{(d)}$ for $d \geq 2$. Then, by (i) and (ii) there ought not to be any efficient result. Indeed, $(F_2^{(d)})_{d \in \mathbb{N}}$, are all fully invariant subgroups of F_2, whence $w \in F_2^{(d)} \subseteq F_2^{(2)}$ implies that any homomorphic image must also belong to $F_2^{(2)}$. In particular, if $\phi = \phi(x, y)$ is such an image, then $\phi \in F_2^{(2)}(x, y) = (X^Y \cap Y^X)' \subseteq (X^Y)' \cap (Y^X)'$ and so, indeed, ϕ is not efficient.

Proof of the Theorem. (i) First we show that (ii) follows from (i). For suppose that w has an efficient result ϕ. Then clearly any group satisfying the law $w \equiv 1$ of length N satisfies the efficient law $\phi(c, d) \equiv 1$, whose length is at most $N \max_i(\ell_{\{c,d\}}(\alpha(x_i)))$. Hence the proof of (ii) will follow directly from the proof of (i).

Accordingly, let $w = w(\mathscr{X})$ be an efficient word. Thus $\mathscr{X} = \{x, y\}$ and $w \in F_2'$.

Part of the method of proof is modeled closely on methods of Semple and Shalev (see [8, 9]), in proving that N-collapsing finite groups are "strongly-locally-nilpotent-by-N-bounded-exponent." We subsequently invoke the aforementioned result of [3].

(1) *Bounding Simple Non-abelian Sections*

Let G satisfy the efficient law

$$w(x, y) \equiv x^{a_1} y^{b_1} x^{a_2} y^{b_2} \cdots x^{a_n} y^{b_n}.$$

We refer to forms (1)–(4) as earlier.

LEMMA 1. *Let G be a finite group satisfying the efficient law $w \equiv 1$ of length N. Then there exists an integer $g = g(N)$ so that G has a soluble fully invariant subgroup H with* $\exp(G/H) \leq g$.

In the course of the proof we shall see what motivates the condition of efficiency.

Proof. Indeed, let $M := A/B \cong \times_{i=1}^{\ell} S_i$ be a non-abelian chief factor of G, where the $(S_i)_{i=1}^{\ell}$ are isomorphic copies of a fixed non-abelian simple group S. Our first aim is to bound $|S|$, the order of S. Recall the result of Jones [4] that a variety of groups generated by infinitely many finite non-abelian simple groups is the variety of all groups.

Now let $\tilde{w}_N$ be the left normed commutator of all possible words w_N of length N in $\{x, y\}$ in a certain specified fixed order, where if we include a word we omit its inverse. Then clearly $\mathscr{V}_{\tilde{w}_N}$, the variety of finite groups satisfying the group law $\tilde{w}_N$, is a proper variety, since $\tilde{w}_N$ is a nontrivial word. Moreover, $\mathscr{V}_{\tilde{w}_N} \supseteq \mathscr{V}_{w_N}$ for every word w_N on $\{x, y\}$ of length N. We thus deduce the existence of a function $h \equiv h(N)$ of N alone, so that if G is a finite group satisfying the law $w \equiv 1$ of length N, and S is a simple non-abelian group involved in G, then $|S| \leq h(N)$.

Now we wish to bound ℓ, the number of direct factors occurring in A/B.

Indeed, let p be a prime dividing $|S|$. The efficiency condition implies that $GCD(\{t_i\}_{1 \leq i \leq m}, p) = 1$ or $GCD(\{v_j\}_{1 \leq j \leq r}, p) = 1$ (or both). Assume, without loss of generality, that the first option holds, and that $p \nmid t_1$, say.

Now consider form (3) for w, and pick an element $x \in G$ whose image (mod B) is contained in S_1, the first copy of S, and has order p modulo B. G acts transitively on the ℓ copies of S by conjugation.

Thus, in particular, if $y \in G$ and $x \in S_1$, then y acts on $x (\operatorname{mod} B)$ by sending x to x^y (mod B), where x^y is an element of order p (modulo B) which is contained in a copy S_i of S, and x^{y^j} corresponds similarly to the jth convolution of y on $x \in S_1$. Our aim is to bound the orbits of elements of G on the copies of S. Indeed, we shall show that these orbits have N-bounded lengths. Accordingly, pick such a y for which $y \notin N_{G/B}(S_1/B)$, the normalizer of S_1/B in G/B. Recall that we have $w(x, y) \equiv 1$, and further, since elements in distinct copies of S commute with each other, we see that

$$w(x,y) \equiv 1 \quad \Rightarrow \quad w(x,y) \equiv 1 (\mathrm{mod}\, B)$$
$$\Rightarrow \quad (x^{t_1})(x^{t_2})^{y^{s_2}} \cdots (x^{t_m})^{y^{s_m}} \equiv 1 (\mathrm{mod}\, B),$$

where we have invoked form (3) for w. Recall also that the exponents s_i of the conjugators y are, by construction, mutually distinct.

But our assumption earlier that p does not divide t_1, implies that $x^{t_1} \not\equiv 1$ (mod B). This then yields a nontrivial identity modulo B between elements in direct factors corresponding to the respective actions of

$$y^0, y^{s_2}, \dots, y^{s_m}.$$

Our conclusion then is that such a y has an orbit on A/B of N-bounded length; explicitly of length at most $N-2$. For if y^{s_k} acts as y^{s_j} for $j<k$, then the orbit has length at most $|s_k - s_j| = |-\sum_{i=1}^{k-1} b_i + \sum_{i=1}^{j-1} b_i| = |-\sum_{i=j}^{k-1} b_i| \leq N-2$, the last inequality following from the fact that w is homogeneous on two letters, and so we must have at least two incidences of x in any such word, thus taking up at least length 2 of our length N word, w.

Since the arguments bounding the orbit of y used only the fixed structure of w_N, we deduce that for any non-abelian chief factor $A/B \simeq \times_{i=1}^{\ell} S_i$, and for any $y \in G$, we have $y^{(N-2)!} \in N_G(S/B)$.

Further, since we already showed that $|S| \leq h(N)$, and in view of the fact that every finite non-abelian simple group S can be generated by two elements (see, for example, [1, Theorem B]), we deduce that $|Aut\, S| \,|\, (h(N)!)^2$.

The net result, in view of all of the above, is that for all $y \in G$, and for all nonabelian chief factors A/B of G, we have

$$y^{(N-2)!(h(N)!)^2} \in C_{G/B}(A/B), \quad \text{i.e.} \quad G^{(N-2)!(h(N)!)^2} \subseteq C_{G/B}(A/B),$$

where $C_{G/B}(A/B)$ is the centralizer of A/B in G/B. But note that

$$\bigcap_{\left(\substack{A/B \text{ perfect} \\ \text{chief factor of } G}\right)} C_{G/B}(A/B)$$

is a *soluble* subgroup of G, and we obtain the desired result, with $H := G^{(N-2)!(h(N)!)^2}$ a *fully invariant* soluble subgroup of G, and

$$\exp(G/H) \leq (N-2)!((h(N))!)^2.$$

∎

(2) *Bounding the Orders of Inner Automorphisms on Abelian Chief Factors*

In view of Lemma 1, and our general aim to show that if G is a finite group satisfying $w \equiv 1$, then G is nilpotent of bounded class by bounded exponent, we may as well assume that G is soluble, and prove:

LEMMA 2. *Let G be a finite soluble group satisfying the group law $w \equiv 1$, where w is as defined earlier. Then there exists a constant $h := h(N)$ so that $G^{h!}$ is a fully invariant nilpotent subgroup of G.*

Proof. Since for any finite group G we have $\operatorname{Fit} G = \bigcap_V C_G(V)$, where V runs over all chief factors of G, and a soluble group has all its chief factors abelian, it is sufficient to prove the following claim:

Claim. Let G be a soluble finite group satisfying the group law $w \equiv 1$. Then there exists an $h := h(N)$ so that $G^{h!} \subseteq C_G(V)$, where V is any (abelian) chief factor of G.

Accordingly, fix such a chief factor $V \cong A/B$, and its characteristic p. Considering, as in the proof of Lemma 1, form (3) or (4) for w according to $GCD(p, \{t_i\}_{1 \leq i \leq m}) = 1$ or $GCD(p, \{v_j\}_{1 \leq j \leq r}) = 1$, without loss of generality let us assume that $GCD(p, \{t_i\}_{1 \leq i \leq m}) = 1$, and thus, choosing any $x \in G$ so that $xB \in A/B$, and $y \in G$, we have

$$w(x,y) \equiv 1 \quad \Leftrightarrow \quad w(x,y) \equiv 1 \pmod{B} \quad \text{or}$$
$$(x^{t_1})(x^{t_2})^{y^{s_2}} \cdots (x^{t_m})^{y^{s_m}} \equiv 1 \pmod{B}, \tag{$*$}$$

where the last equality is obtained by invoking form (3) for w.

Since V is an abelian section, of characteristic prime to $GCD_{1 \leq i \leq m}\{t_i\}$, this implies that every $y \in G \setminus C_G(V)$ is a zero of the "Laurent" polynomial $f(\lambda) \equiv t_1 + t_2\lambda^{s_2} + t_3\lambda^{s_3} + \cdots + t_m\lambda^{s_m}$ in its action on V. Note that we may replace $f(\lambda)$ by a monic polynomial in the following manner.

Since $w(x,y) \equiv 1 \ \Leftrightarrow \ y^{\alpha}w(x,y)y^{-\alpha} \equiv 1$ for any integral power α, we have $f(\lambda) \equiv 0 \ \Leftrightarrow \ \lambda^{-\alpha}f(\lambda) \equiv 0$ in their respective actions on V. Now choose $\alpha = \min((s_i)_{1 \leq i \leq m}, 0)$ and $\lambda^{-\alpha}f(\lambda)$ is thus such a polynomial.

Since this polynomial acts on a vector space of characteristic p, we may reduce each t_i modulo p without changing the effect of the action. The fact that $GCD(p, \{t_i\}_{1 \leq i \leq m}) = 1$ guarantees that the polynomial is nontrivial, and, as can be seen by applying considerations similar to those in the proof of Lemma 1, is of degree at most $N - 2$. Denote the polynomial thus obtained by $f_p(t)$. Let $\operatorname{Ann}(V)$ be the annihilator of V. Clearly, if $y \in G \setminus C_G(V)$, then y and all its integral powers are zeros of $f_p(t)$ in their actions on V. In particular, we have $f_p(y), f_p(y^2) \ldots, f_p(y^{N-1}) \in \operatorname{Ann}(V)$.

We now apply the following variation of a ring-theoretic lemma from [8, 9] (see [8, Lemmas 3.3–3.4], [9, Lemma 3.1]).

LEMMA. *Let $f_p(t)$ be as above, with I the ideal generated by*

$$f_p(t), f_p(t^2), \ldots, f_p(t^{N-1}) \text{ in } F_p[t].$$

Then there exist positive integers k and m so that $(t^m - 1)^k \in I$, with k and m functions of N only.

So in particular, since $I \subseteq \mathrm{Ann}(V)$, we see that the identity $[V, {}_k G^m] \equiv 1$ holds in G. Applying techniques from [9, Lemma 3.2], we deduce further that there exists a constant $e = e(k, m)$ so that if $w(x, y)$ is an efficient law in G, and V is as above, an elementary abelian p-section of G, then $y^e - 1 \in \mathrm{Ann}(V)$, or equivalently, that the identity $[V, G^e] \equiv 1$ holds in G; or $G^e \subseteq C_G(V)$.

But now, since there are an N-bounded number of possible efficient words of length N, and each corresponding polynomial is derived from such a word and is of N-bounded degree, we thus deduce that k, m, and so e are bounded as functions of N alone, and

$$G^e \subseteq \bigcap_{V \text{ chief factor of } G} C_G(V) = \mathrm{Fit}\, G \quad \text{is nilpotent.}$$

∎

(3) *Bounding the Nilpotency Class*

In the previous two lemmas we showed that if G is a finite group satisfying the efficient law $w_N \equiv 1$ of length N, then there is an $e = e(N)$ so that G^e centralizes every chief factor of G, and is, in particular, nilpotent. At this point we remark that the considerations employed in the proof of Lemma 2 leading up to the lemma following it imply that the hypothesis that G, a soluble group, satisfies an efficient law of length N in $\{x, y\}$, yields the same results as the assumption that G is n-collapsing (with bounds functions of N instead of n).

Accordingly, the results in [9, Sects. 4 and 5] yield the following analogue to Theorem B therein;

LEMMA 3. *There exist functions f, g such that every finite group G satisfying the efficient law $w_N(x, y)$ of length N in $\{x, y\}$ possesses a nilpotent normal subgroup K, satisfying*

(1) $\exp(G/K)$ *divides $f(N)$, and*

(2) *Every d-generated subgroup of K has nilpotency class at most $g(N, d)$.*

All that now remains is to eliminate the dependence of the nilpotency class of $G^{f(N)}$ on d, the number of generators of any nilpotent subgroup.

Indeed, Lemma 3 implies that $\mathscr{B}$, the variety generated by all finite groups satisfying the efficient law $w \equiv 1$, is locally nilpotent-by-finite. Thus we may invoke the following result of Burns et al. [3].

THEOREM C. *A variety $\mathscr{B}$ of groups is locally nilpotent-by-finite* (*i.e.*, *has all of its finitely generated subgroups nilpotent-by-finite*) *if and only if $\mathscr{B} \subseteq \mathscr{N}_c B_e$ for some c, e.*

This completes the proof of parts (i) and (ii) of the theorem. ∎

We now turn to

Proof (iii). (1) First, note that if G is a finite group of bounded exponent, say $G^e = 1$, then G satisfies the law $[x, y^e] = 1$, which is clearly efficient and of e-bounded length.

(2) In the general case, let us assume that G is nilpotent-of-class c-by-exponent-c. Groups this type are known to satisfy a *positive* law of c-bounded length (see [6, 9] for proofs).

(3) Without loss of generality then, we may assume that the positive law h assumed in (2) is a law on two letters (actually, the concrete construction of such a law is demonstrated in [9]), and that $h = h(x, y)$ is *homogeneous*. (For otherwise make the substitution $x := \alpha$, $y := \alpha$, rendering G a group of bounded exponent, which has been dealt with in (1).)

(4) We shall see in Section II (see Corollary 1 therein) that every homogeneous word yielding a positive efficient law of length N has a result of N-bounded length which is an efficient word.

Thus, modulo the proof of statement (4), we have outlined the proof of part (iii) of the theorem. ∎

SECTION II

In the previous section we learned that varieties generated by nilpotent-of-bounded-class-by-bounded-exponent finite groups are characterized by their satisfying an efficient law. Conversely, a word defines such a variety if it has an efficient result.

We thus need to tackle the following questions:

(i) Which types of words have an efficient result?

(ii) Which types of words do *not* have an efficient result?

We shall tackle the first question by outlining some techniques which explicitly construct efficient results in some cases, and, in view of the theorem, the second question can be answered by displaying families of finite groups which on the one hand satisfy the word in question and on the other hand are not of uniformly bounded nilpotency classes by uniformly bounded exponent.

We demonstrate the rudiments of these techniques with some examples;

EXAMPLES. (1) Consider $w_{3,4} := [x^4, y^4]$ from the previous section, which we noted to be a word which is not efficient. Apply the substitution $x \mapsto b$, $y \mapsto ba$ to yield

$$\begin{aligned}\phi(a,b) &= b^{-4}(ba)^{-4}b^4(ba)^4\\ &= b^{-4}a^{-1}b^{-1}a^{-1}b^{-1}a^{-1}b^{-1}a^{-1}b^{-1}b^4 babababa\\ &= (a^{-1})^{b^4}(a^{-1})^{b^5}(a^{-1})^{b^6}(a^{-1})^{b^7}a^{b^3}a^{b^2}a^b a.\end{aligned}$$

Here we have expressed ϕ in form (1), but since all conjugating powers for b are distinct, this is also form (3). Since all the exponents of the a's are ± 1, we already see that ϕ is an efficient result of w without considering forms (2) and (4).

(2) Similar techniques can be applied (see the proposition in this section for a proof and Example 3 following it) to show that the family, $w_{3,n_0} = [x^{n_0}, y^{n_0}]$ always has an efficient result of n_0-bounded length and thus defines a variety of finite groups which are nilpotent of n_0-bounded class by n_0-bounded exponent. (In fact it is clear that G satisfies $w_{3,n_0} \Leftrightarrow \gamma_2(G^{n_0}) \equiv 1$.)

(3) On the other hand, we claim that the family $w_{4,n_0} = [x, y]^{n_0}$ does not have an efficient result for any $n_0 > 1$!

Proof. By the theorem, if w_{4,n_0} had an efficient result, it would define a variety $V_{w_{4,n_0}}$ which would be nilpotent of bounded class by bounded exponent. On the other hand, if p is any prime dividing n_0, then the family $G_n := \{C_p wr C_n\}_{n\to\infty}$ satisfies w_{4,n_0} and is clearly *not* contained in a product of varieties $\mathcal{N}_{f+1}B_f$ described earlier. ∎

We now generalize the phenomenon seen in Example 1 above, in the following:

PROPOSITION. *Let $w(x, y) \equiv x^{a_1}y^{b_1}x^{a_2}y^{b_2}\cdots x^{a_n}y^{b_n}$ be a homogeneous word of length N.*

Consider the lists $(a_1, a_2, a_3, \ldots, a_n)$ (resp. $(b_1, b_2, b_3, \ldots, b_n)$) *as "cycles" of length n, by which we mean that a_2 follows a_1, a_3 follows $a_2, \ldots,$ and a_1 follows a_n in cyclical order (and similarly, that b_2 follows $b_1, \ldots, b_1$ follows*

b_n). *Suppose that there exists an* a_i *such that there is no proper subset* $a_{i+1}, a_{i+2}, \dots, a_s$ *of consecutive* a_j*'s following* a_i, *so that* $\sum_{j=i+1}^{s} a_j = 0$, i.e., *there is no* s, $i+2 \le s \le i-1$ *with* $\sum_{j=i+1}^{s} a_j = 0$, (*or similarly that there exists a* b_k *with no* m, $k+2 \le m \le k-1$ *so that* $\sum_{\ell=k+1}^{m} b_\ell = 0$).

Then w *has an efficient result of* N*-bounded length, whence* w *defines a nilpotent-of-bounded-class-by-bounded-exponent variety generated by finite groups, whose bounds are functions of* N *alone.*

Note. The condition above, which we shall call the "cyclical condition," means that there exists an a_i (or a b_k), for which we cannot find a chain of *subsequent consecutive* a_i's (or b_j's) summing to zero, without using all the a_i's in the word (in which case we *always* get zero sum by homogeneity). In this case we say that a_i (or b_k) *supplies the cyclical condition.*

EXAMPLES. (1) Consider the following word $w(x, y) \equiv [x^2, y^2, x^2]$. Clearly w is not efficient, yet we shall see that w satisfies the cyclical condition (c.c.). Indeed,

$$w(x, y) \equiv y^{-2}x^{-2}y^2x^{-2}y^{-2}x^2y^2x^2.$$

Consider the corresponding list

$$\{b_1, a_2, b_2, a_3, b_3, a_4, b_4, a_1\} := \{-2, -2, 2, -2, -2, 2, 2, 2\}.$$

Checking each a_i, b_i in turn, we note that $b_2 + b_3 = 0$, so b_1 does not supply the c.c. similarly, $a_3 + a_4 = 0$, so a_2 is "no good," nor is b_2, as $b_3 + b_4 = 0$. Yet a_3 supplies the cyclical condition, as $a_4 + a_1 = 4$, $a_4 + a_1 + a_2 = 2$, and $a_4 + a_1 + a_2 + a_3 = 0$, so we only reach zero sum by taking all the a_i's in the word.

In view of the proposition, w is guaranteed to have an efficient result; indeed, substituting $x \mapsto b$, $y \mapsto ba$ yields

$$w(x, y) \equiv \phi(a, b) \equiv (a^{-1})(a^{-1})^{b}(a^2)^{b^2}(a^2)^{b^3}(a^{-1})^{b^4}(a^{-1})^{b^5}c,$$

and ϕ is an efficient word.

(2) Similarly, any word $w_{n_0}(x, y) \equiv [x^{n_0}, y^{n_0}, x^{n_0}]$ satisfies the c.c., since the "pattern" of the cycle is the same as that in Example 1. Thus w_{n_0} has an efficient result of n_0-bounded length, and so $\mathscr{V}_{w_{n_0}} \subseteq \mathscr{N}_{f(n_0)}\mathscr{B}_{f(n_0)}$ for some function $f(n_0)$ of n_0.

(3) The examples $w_{3,n_0}(x, y) = [x^{n_0}, y^{n_0}] = x^{-n_0}y^{-n_0}x^{n_0}y^{n_0}$ also satisfy the c.c., hence they have an efficient result.

(4) Returning to the family $w_{4,n_0} = [x, y]^{n_0}$ for $n_0 > 1$, it is clear that they do not satisfy the c.c. (every a_i and every b_k have $n_0 - 1$ proper subsums to zero)! This is in keeping with the counterexamples produced earlier for this family of words.

(5) In the same vein, it can be checked quite easily that any commutator of the form $[[\alpha, \beta], [\gamma, \delta]] \in F^{(2)}$, with $\alpha, \beta, \gamma, \delta$ any words in $\{x, y\}$, does not satisfy the c.c.; hence, as is easily seen, nor does any product of such commutators.

We now turn to the proof of the proposition, which also supplies an effective method for finding an appropriate substitution for words satisfying the c.c. such as w in Example 1 following the proposition.

Proof of the Proposition. Let $w(x, y)$ be of length N and satisfy the c.c. Our aim is to show that w has an efficient result of N-bounded length. We shall do so by concretely constructing an appropriate substitution $x \mapsto \alpha(c, d)$, $y \mapsto \beta(c, d)$ so that $\phi(c, d) := w(\alpha, \beta)$ is an efficient word.

So let $x \mapsto d^{\ell}c$, $y \mapsto d$ in $w(x, y)$, where our hope is to determine ℓ at a later stage in such a way as to render ϕ efficient. It turns out, as we shall see in the process of the proof, that the c.c. is a condition which makes this choice possible. Applying the above substitution to w yields

$$w(d^{\ell}c, d) = (d^{\ell}c)^{a_1} d^{b_1} (d^{\ell}c)^{a_2} d^{b_2} \cdots (d^{\ell}c)^{a_n} d^{b_n}. \qquad (*)$$

We now examine the "contents" of the bracket $(d^{\ell}c)^{a_1}$ obtained when expressing it as a product of conjugates of c by powers of d, giving

$$(d^{\ell}c)^{a_1} = \begin{cases} c^{d^{-\ell}} c^{d^{-2\ell}} c^{d^{-3\ell}} \cdots c^{d^{-a_1\ell}} d^{a_1\ell} & \text{if } a_1 > 0 \\ c^{-1}(c^{-1})^{d^{\ell}} (c^{-1})^{d^{3\ell}} \cdots (c^{-1})^{d^{-(a_1+1)\ell}} d^{a_1\ell} & \text{if } a_1 < 0. \end{cases}$$

Considering subsequent brackets, note that powers of the form d^{b_i} "contribute" to the powers of d by which subsequent c's or c^{-1}'s are already conjugated.

It is readily checked that one obtains the formula for $\phi(c, d) = w(d^{\ell}c, d)$

$$\phi(c, d) = \prod_{j=1}^{n} \prod_{r=1}^{a_j} (c^{\operatorname{sign}(a_j)})^{d^{m_{j,r}}}, \qquad (**)$$

where

$$m_{j,r} = m_{j,r}(\ell) = \begin{cases} -\left[\sum_{s=1}^{i-1} b_s + \left(r + \sum_{s=1}^{j-1} a_s\right)\ell\right] & \text{if } a_j > 0 \\ -\left[\sum_{s=1}^{j-1} b_s + \left(-r + 1 + \sum_{s=1}^{j-1} a_s\right)\ell\right] & \text{if } a_j < 0. \end{cases}$$

Comparing $\phi(c, d)$ with an efficient $w(x, y)$ in form (3), with c playing the role of x and d the role of y, our aim now is to *fix* ℓ so that "$GCD(\{t_i\}_i)$" $= 1$.

Note that all exponents of c in the word $\phi(c, d)$ in the form $(**)$ above, are equal to ± 1. Thus, if we succeed in fixing $\ell > 0$ that at least one of the $m_{j,r}$'s does not equal any of the others, then we are assured that even after commutator collection of $\phi(c, d)$ to obtain form (3), this will remain the case, thus rendering $GCD(\{t_i\})_{1 \le i \le m}$ equal to 1, and we will have obtained an efficient ϕ.

Choosing an ℓ

A key point to note here is that

$$w(x, y) \equiv x^{a_1} y^{b_1} x^{a_2} \cdots x^{a_n} y^{b_n} \equiv 1 \qquad \Leftrightarrow$$
$$x^{a_i} y^{b_i} \cdots x^{a_n} y^{b_n} x^{a_1} y^{b_1} x^{a_1} b^{b_2} \cdots x^{a_{i-1}} y^{b_{i-1}} \equiv 1 \quad \text{for any } 1 \le i \le n.$$

Thus, since w by hypothesis satisfies the c.c., we may assume without loss of generality that b_1 supplies the c.c. for w (this means that $w(x, y)$ is written in the form $x^{a_1} y^{b_1} \cdots x^{a_n} y^{b_n}$ above and that there is no k, $3 \le k \le n$, with $\sum_{j=2}^{k} b_j = 0$). Consequently, consider the conjugates of c arising out of the bracket $(d^{\ell} c)^{a_2}$ in $(*)$ above.

The corresponding $m_{2,r}$, $1 \le r \le |a_2|$, have values

$$m_{2r}(\ell) = \begin{cases} -b_1 - (r + a_1)\ell & \text{if } a_2 > 0, \\ -b_1 - (1 - r + a_1)\ell & \text{if } a_2 < 0. \end{cases}$$

So, in particular,

$$m_{2,1}(\ell) = \begin{cases} -b_1 - (a_1 + 1)\ell & \text{if } a_2 > 0, \\ -b_1 - a_1 \ell & \text{if } a_2 < 0. \end{cases}$$

We shall fix ℓ so that $m_{21}(\ell) \ne m_{j,r_j}(\ell)$ for any other exponent m_{j,r_j}, $1 \le j \le n$, $1 \le r_j \le |a_j|$. In essence, the m_{j,r_j} are all linear functions in the indeterminate ℓ, for each m_{j,r_j} is an expression of the form

$$m_{j,r_j}(\ell) = \alpha_{j,r_j} + \beta_{j,r_j} \ell.$$

Now solve the linear equation

$$m_{j,r_j}(\ell) = m_{2,1}(\ell). \qquad (***)$$

An integral solution of such an equation $(***)$, if it exists, consists of an ℓ for which, if we "use" it, we will obtain $m_{j,r_j} \equiv m_{2,1}$, thus, a term with a conjugator $d^{m_{21}} = d^{m_{j,r_j}}$ will appear more than once in the word. If so, we risk that after the resulting $\phi(c,d)$ is expressed in form (3) for all ℓ, the exponent of c conjugated by $d^{m_{21}}$ might thus be zero or of absolute value greater than 1. Note that there are $\sum_{j=1}^{n}|a_j| \leq N-2$ such m_{j,r_j}'s.

So we solve all of the $\leq N-3$ such equations of the form $(***)$ and then pick an ℓ so that it is none of these integral solutions. (We may as well shorten the length of the word $\phi(d^{\ell}c, d)$ thus obtained as much as possible, so, choosing ℓ of minimal absolute value, we can be guaranteed to find one of absolute value at most $\lfloor (N-2)/2 \rfloor$). The only case when an ℓ *cannot* be found, is the case when $m_{21}(\ell) \equiv m_{j,r_j}(\ell)$ *identically* for some m_{j,r_j}, namely, when $\alpha_{j,r_j} = \alpha_{2,1}$ *and* $\beta_{j,r_j} = \beta_{21}$.

But the c.c. actually prevents this.

For $\beta_{21} = \beta_{j,r_j} \Leftrightarrow -b_1 = -\sum_{i=1}^{j-1} b_i \Leftrightarrow \sum_{i=2}^{j-1} b_i = 0$, and this does not happen for any j, since b_1 supplies the cyclical condition.

Thus the claim in the proposition has been proved, and thus, if w satisfies the c.c., then w has a p.c. of length at most $\lfloor (N-2)/2 \rfloor \cdot (N-2) + 2$ and hence defines a nilpotent-of-N-bounded-class-by-N-bounded-exponent variety of finite groups.

A WORKED EXAMPLE. In Example (1), following the statement of the proposition, we saw that while the word $w(x,y) = [x^2, y^2, x^2]$ is not efficient, it satisfies the c.c., since $w(x,y) = y^{-2}x^{-2}y^2x^{-2}y^{-2}x^2y^2x^2$. Recall that a_3 supplies the c.c; so let us rewrite the word beginning with b_2 (to correspond to the situation as in the proof), giving $w(x,y) = y^2x^{-2}y^{-2}x^2y^2x^2y^{-2}x^{-2}$. We subsequently substitute $y \mapsto d^{\ell}c$, $x \mapsto d$ (as here the roles of x and y are interchanged versus those in the proof since the c.c. was obtained via an a_i not a b_i). This gives

$$\phi_{\ell}(c,d) = (d^{\ell}cd^{\ell}c)d^{-2}(c^{-1}d^{-\ell}c^{-1}d^{-\ell})d^2(d^{\ell}cd^{\ell}c)d^2(c^{-1}d^{-\ell}c^{-1}d^{-\ell})d^{-2}$$

or

$$\phi_{\ell}(c,d) = c^{d^{-\ell}}c^{d^{-2\ell}}(c^{-1})^{d^{-2\ell+2}}(c^{-1})^{d^{-\ell+2}}c^{d^{-\ell}}c^{d^{-2\ell}}(c^{-1})^{d^{-2\ell-2}}(c^{-1})^{d^{-\ell-2}}.$$

The third element in the list of conjugates of c by powers of d above is the one corresponding to "m_{21}" in the proof.

So we solve the following equations:

	solutions
$-2\ell + 2 = -\ell$	$\ell = 2$
$-2\ell + 2 = -2\ell$	no solution
$-2\ell + 2 = -\ell + 2$	$\ell = 0$
$-2\ell + 2 = -2\ell - 2$	no solution
$-2\ell + 2 = -\ell - 2$	$\ell = 4$

Thus we see that the only ℓ's to avoid are $\ell = 0, 2, 4$. (Note that $\ell = 0$ corresponds to the original word!)

We thus pick $\ell = 1$, a choice of smallest absolute value, and then $\phi(c, d) = (c)^{d^{-1}}(c)^{d^{-2}}(c^{-1})(c^{-1})^{d}(c)^{d^{-1}}(c)^{d^{-2}}(c^{-1})^{d^{-4}}(c^{-1})^{d^{-3}}$, which, written in form (3), has the form

$$\phi_1(c, d) = (c^{-1})^{d^{-4}}(c^{-1})^{d^{-3}}(c^2)^{d^{-2}}(c^2)^{d^{-1}}(c^{-1})(c^{-1})^{d}\big(\mathrm{mod}(C^D)'\big),$$

which is efficient (where, as usual $C = \langle c \rangle$, $D = \langle d \rangle$ in $F(c, d)$). ∎

Note. (i) Comparing the Laurent polynomial associated with ϕ above to that with ϕ in Example (1) following the statement of the proposition, we see that they differ only by a factor λ^4. This difference is due to the fact that in obtaining ϕ here, we applied a "cyclical shift" first.

(ii) In effectively finding a substitution, clearly any $m_{s,t} = \alpha + \beta\ell$ which is not *identically* equal to any other $m_{r,v}$ will do in place of $m_{2,1}$.

We are now in a situation to state and prove Corollary 1, thereby completing the proof of the theorem in the previous section.

COROLLARY 1. *Let $\mathscr{X} = \{x, y\}$, and let $w = w(\mathscr{X})$ be a homogeneous word of length N in $\mathscr{X}$ which yields a positive law. Then w has an efficient result of N-bounded length.*

Proof. By the proposition, we shall simply show that all such laws satisfy the c.c. But this is evident. Write $w(x, y)$ in the form uv^{-1}, where

$$u(x, y) = x^{a_1}y^{b_1}x^{a_2} \cdots x^{a_n}y^{b_n}$$

and

$$v(x, y) = y^{c_1}x^{d_1} \cdots y^{c_m}x^{d_m}$$

with all exponents positive, and $\sum_{i=1}^{n} a_i = \sum_{j=1}^{m} d_j$ and $\sum_{i=1}^{n} b_i = \sum_{j=1}^{m} c_j$, and again in the form

$$w(x, y) = x^{a_1}y^{b_1} \cdots x^{a_n}y^{b_n}x^{-d_m}y^{-c_m} \cdots x^{-d_1}y^{-c_1}.$$

We see that clearly the exponent $-d_1$ supplies the c.c. For, since all the exponents a_i for $i = 1, \dots, n$ are positive, say, if $\sum_{i=1}^{n} a_i = N$, then, since all the exponents $-d_j$ for $j = 1, \ldots, m$ are negative, and the whole sum $\sum_{i=1}^{n} a_i - d_1 - d_2 - \cdots - d_m$ equals zero, one cannot possibly obtain a total sum of zero without using *all* the terms $a_1, \ldots, a_n$, and $-d_m, \ldots, -d_1$. ∎

To conclude this section, we mention the following corollary, whose proof is based on ideas similar to those in Corollary 1, and whose details we omit.

COROLLARY 2. *Let $w_N(x, y)$ be a homogeneous word of length N. Suppose there exists a b_i (or, in what follows, one may equally state the corresponding conditions for an a_j), so that, if*

$$\sum_{j=i+1}^{s_1} b_j = \sum_{j=i+1}^{s_2} b_j = \sum_{j=i+1}^{s_3} b_j = \cdots = \sum_{j=i+1}^{s_r} b_j = 0,$$

are all the subsums summing to zero following b_i, then we have $|a_{i+1}| > |a_{s_1+1}| + |a_{s_2+1}| + \cdots |a_{s_r+1}|$. Then w has an efficient result of N-bounded length.

We now refer to the example mentioned in the Introduction.

EXAMPLE.

$$w(x, y) \equiv x^{-2}y^{-2}x^{2}y^{2}x^{-10}y^{-2}x^{10}y^{2}x^{-2}y^{-4}x^{-6}y^{6}x^{-2}y^{-2}x^{10}y^{-4}x^{-8}y^{6}x^{8}y^{-2}.$$

It can be checked that w above is not efficient, does not satisfy the cyclical condition, but *does* satisfy the hypotheses of Corollary 2; hence w defines a variety generated by nilpotent-of-bounded-class-by-bounded-exponent finite groups!

SECTION III

In this short section we aim to give an indication of some useful consequences which can be derived from the theory developed in the previous sections. It turns out that if we are prepared to restrict the set of finite groups somewhat, we can obtain corresponding results which are true for a much wider class of words.

DEFINITION. Call a finite group *π-free* if its order is not divisible by any member of the set π. For any $N \in \mathbb{N}$, denote by π_N the set of primes p with $p < N$.

PROPOSITION 1. *Let $w = w_N(x, y)$ be* any *word of length N on an alphabet $\mathscr{X} = \{x, y\}$ such that $w \notin (X^Y)' \cap (Y^X)'$. Then the variety generated by all π_N-free finite groups satisfying w_N is contained in $\mathscr{N}_{c+1}\mathscr{B}_c$ for some c which is N-bounded.*

Proof. Clearly we may assume that w is homogeneous. Thus, considering $w(x, y)$ in form (3), we have

$$w(x, y) \equiv x^{t_1}(x^{t_2})^{y^{s_2}} \cdots (x^{t_m})^{y^{s_m}} c, \quad \text{where } c \in (X^Y)',$$

and, since w has length N and is homogeneous, we must have $|t_i| < N$ for all $1 \leq i \leq m$.

Now let G be any π_N-free finite group satisfying $w_N \equiv 1$. Then it is evident that we may successfully consider w_N in form (3) above throughout the proof of part (i) of the theorem in Section I, to obtain the same results.

Notes. (1) We may loosely summarize Proposition 1 as saying "w is efficient for π_N-free groups."

(2) Clearly, concrete examples can yield corresponding results for varieties less restricted than π_N-free varieties, as the following example demonstrates:

EXAMPLE. Consider $w(x, y) \equiv [x^3, y^3]^{10}$.

As we discussed earlier, $\mathscr{V}_w$ cannot be contained in the product of a nilpotent and a locally finite Burnside variety, as it contains groups of the forms $C_2 wr C_n$ and $C_5 wr C_m$ for n and m as large as we please.

Applying Proposition 1, since w is of length 120, we may conclude that the variety generated by finite groups, none of whose orders are multiples of any prime less than 120, is contained in a product of varieties of the form $\mathscr{N}_{c+1}\mathscr{B}_c$ for bounded c.

However, one can improve on this, for, noting that $[x^3, y^3]$ satisfies the cyclical condition, a substitution such as $x \mapsto dc$, $y \mapsto c$ renders $[x^3, y^3]$ efficient; thus $w(dc, c)$ is a result which, when written in form (3) yields a word for which (refer to notation in Sect. I) $GCD(\{t_i\}_{1 \leq i \leq m}, \{v_j\}_{1 \leq j \leq r}) = 10$. Thus we see that the only "obstructions" to efficiency, that is, the only primes whose section w *does not* control with respect to the above substitution and *cannot* control due to the counterexamples mentioned above, are 2 and 5.

Thus, we may conclude that the variety generated by $\{2, 5\}$-free finite groups satisfying w is nilpotent-of-bounded-class-by-bounded-exponent.

Proposition 1 enables us to derive a variant of part (ii) of the theorem using weaker hypotheses.

COROLLARY. *Let w be as in Proposition 1, thus $w = w_N(x, y) = x^{a_1} y^{b_1} \cdots x^{a_n} y^{b_n}$, a homogeneous word of length N, with $w \notin (X^Y)' \cap (Y^X)'$.*

Suppose that for every prime p, w has a result w_p whose length is p-bounded, and for which (using previous notation), $GCD(\{t_i\}_{1 \leq i \leq m}, \{v_j\}_{1 \leq j \leq r}, p) = 1$.

Then w defines a variety generated by finite groups which are nilpotent-of-class-at-most-f-by-exponent-at-most-f, where f is a function of N alone.

Note, first, that the corollary is a generalization of part (ii) of the theorem.

The difficulty at first glance with the hypotheses in the corollary is that we do not know the lengths of $(w_p)_{p \text{ prime}}$ to be bounded *uniformly* for all p.

Proof of the Corollary. We shall show the corollary to be true by constructing an efficient word of N-bounded length which is a law whenever w is, and then we are in the situation of part (i) of the theorem.

Denote, for any word w, $GCD_w(\{t_i\}_{1 \le i \le n}, \{v_j\}_{1 \le j \le r})$ to be as previously described, but referred to w.

In Proposition 1, we saw that if $p > N$, then, considering form (3) for w, we have $GCD(\{t_i\}_{1 \le i \le m}, p) = 1$.

Since the set π_N is finite, so clearly is the set $\{w_p\}_{p \in \pi_N}$. For each such p, the hypothesis $GCD_{w_p}(\{t_i\}_{1 \le i \le m}, \{v_j\}_{i \le j \le r}, p) = 1$ clearly implies that either $GCD_{w_p}(\{t_i\}_{1 \le i \le m}, p) = 1$ or $GCD_{w_p}(\{v_j\}_{1 \le j \le r}, p) = 1$. In the event that the second alternative holds, replace $w_p(x, y)$ by the equivalent word $w_p(y, x)$.

Now for any word w, $w = 1 \Leftrightarrow w^{y^k} \equiv 1$ for any integer k (where w^{y^k} denotes, as usual, conjugation of w by y^k). Now pick any $\ell \in \mathbb{N}$ with $\ell > \max(\text{length}(w_p)_{p \in \pi_N}, N)$, and consider the word

$$W(x, y) = w \cdot \prod_{i=1}^{t} w_{p_i}^{y^{\ell i}}, \quad \text{where } \pi_N = \{p_1, \dots, p_t\}.$$

Then W is a law whenever w is, and further, the conjugation of the words w_{p_i}, for $p_i \in \pi_N$, by successively higher powers of y^ℓ has the effect that conjugates of x by powers of y in different words w_p cannot "double up" or "cancel out" with each other in the word W. The net effect of the construction is that $GCD_W(\{t_i\}_{1 \le i \le m}) = 1$ and W is an efficient word.

Note that its length is a function of N alone, since each w_{p_i} has N-bounded length, and there are an N-bounded number of them.

SECTION IV. OPEN QUESTIONS

(1) We have seen that the word $[x, y]^{n_0}$ has no efficient result for any $n_0 > 1$. The same technique yields that $w \equiv v^{n_0}$, where v is any homogeneous word on two letters, has no efficient result. Suppose $w(x, y) = v_1^{n_1} v_2^{n_2} \cdots v_r^{n_r}$ is a homogeneous word with $n_i > 1 \quad 1 \le i \le r$,

$GCD(\{n_i\}_{1\leq i\leq r}) = 1$, and v_i, $1 \leq i \leq r$ all have efficient results. Does that imply that w has an efficient result?

(2) If $w(x, y) \in (X^Y)' \cap (Y^X)'$ does that imply that w has no efficient result? (Note that the following example due to Professor B. Plotkin shows that $F_2(x, y)^{(2)} \subsetneqq (X^Y)' \cap (Y^X)'$. Indeed it may be checked that the element $u := [[[x, y], x], y]$ in $F_2(x, y)$ belongs to $(X^Y)' \cap (Y^X)'$ but is not an element in $F_2(x, y)^{(2)}$.)

(3) A similar study could be embarked or. to determine "which words spell almost soluble." Two partial results can be given in this direction but it would be interesting to be able to solve the problem in more generality.

(i) First, let n_0 be any power which is not a multiple of $\exp(S)$, for any simple non-abelian group S. Then for any homogeneous word w with an efficient result, it follows from results of this chapter that w^{n_0} defines a variety of finite groups which are soluble by n_0-bounded exponent. For example, in particular, by the Odd Order Theorem of Feit and Thompson, if n_0 is odd. However, it would be interesting to determine whether the solubility lengths are bounded for such words.

(ii) Second, if we consider any proper variety generated by finite groups of ranks at most r, it has been shown (see [2]) that within that context, *all* words spell almost soluble, even almost nilpotent-by-abelian-by-bounded-exponent, with bounds for nilpotency class and exponent being functions of the length of the word and the rank r alone.

ACKNOWLEDGMENTS

I thank my supervisor, Professor Alexander Lubotzky, for his guidance during the preparation of this work which formed part of my doctoral thesis. I am grateful to the referee for suggestions which improved the exposition of the paper.

REFERENCES

1. M. Aschbacher and R. Guralnick, Some applications of the first cohomology group, *J. Algebra* **90** (1984), 446–460.
2. S. Black, A Finitary Tits Alternative, *Arch. Math.* **72** (1999), 86–91.
3. R. Burns, O. Macedonska, and U. Medvedev, Groups satisfying semigroup laws, and nilpotent by Burnside varieties, *J. Algebra* **195** (1999), 510–525.
4. G. A. Jones, Varieties and simple groups, *J. Austral. Math. Soc.* **17** (1974), 163–173.
5. Y. K. Kim and A. H. Rhemtulla, Weak maximality condition and polycyclic groups, *Proc. Amer. Math. Soc.* **123**, No. 3 (1995), 711–714.
6. A. I. Malcev, Nilpotent Semigroups, *Uchen. Zap. Ivanovsk. Ped. Inst.* **4** (1953), 107–111.
7. F. Point, Milnor Identities, *Commun. Algebra* **24**, No. 12 (1996), 3725–3745.

8. J. F. Semple and A. Shalev, Combinatorical conditions in residually finite groups I, *J. Algebra* **157** (1993), 43–50.
9. A. Shalev, Combinatorical Conditions in residually finite groups II, *J. Algebra* **157** (1993), 51–62.
10. J. S. Wilson, Two generator conditions for residually finite groups, *Bull. London Math. Soc.* **104** (1991), 239–248.
11. E. I. Zelmanov, The solution of restricted Burnside problem for groups of prime power exponent, Yale University Notes (1990).
12. E. I. Zelmanov, On some problems of group theory and Lie algebras, *Math. USSR-Sb*, **66** (1990), 159–168.
13. E. I. Zelmanov, ''Nil Rings and Periodic Groups,'' KMS Lecture Notes in Mathematics, Korean Math. Soc., Seoul, 1992.